

2020-04-24

Labels strengthen motor learning of new tools

Foerster, FR

<http://hdl.handle.net/10026.1/15601>

10.1016/j.cortex.2020.04.006

Cortex

Elsevier BV

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

Labels strengthen motor learning of new tools

Abstract

Recent reformulations of the Sapir-Whorf hypothesis have shown how labels can guide our thinking in situations of uncertainty, facilitating the identification of objects. We examine whether the effect of labels extends beyond perceptual processes, to help us learn the motoric manipulations required to use novel tools. Exploiting immersive virtual reality, we measured behavioural movement latencies and electrophysiological activity from participants learning to use a range of labeled and unlabeled novel tools. We found that providing a tool with a label reduced the time taken to reach for it, with participants also faster and more accurate when executing the manipulations required to use it. Conversely, labels did not confer any facilitation when the tool was simply moved to another location; participants were slower to grasp a labeled tool when asked to transport it. These findings were also supported by electrophysiological recordings, showing a reduction in sensorimotor beta-band (~30Hz) power when participants were asked to use the labeled tools, but not move them. This modulation of beta activity indicates augmented learning of motor-activity related to tool use within somatosensory regions due to the activation of its lexical representation. These results suggest an extension of the Whorfian hypothesis, such that language not only modulates our thoughts and perceptual processes, but also affects our actions with objects and tools. We propose that labels tune our somatosensory experience and help to memorize body states related to tool use by creating an invariant lexical anchor on which we can build motor learning and experience.

Significance Statement

Language and skilled tool use are pivotal endowments of our species. We investigated the role of tool labels in motor learning and brought behavioural evidence that labels help to learn, remember and perform tool use. This enhancement is reflected in the beta-band power, reflecting augmented sensorimotor processing induced by the labels. The results explain why humans attribute labels to tools: labels not only allow us to communicate about tools and help us to identify their referents, they also enhance their usage. This finding extends the Sapir-Whorf hypothesis that language influences the way we think, to the way we act and blurs the distinction between linguistic and motor processing.

Keywords: label; motor learning; tool use; action selection

Labels strengthen motor learning of new tools

1. Introduction

Language and tool use are endowments of our species, sharing evolutionary origins (Arbib, 2011) and neuroanatomical implementations (Frey, 2008). Our different spoken language influences perception in a variety of domains, such as those related to the perception of color (Regier and Kay, 2009; Thierry et al., 2009; Winawer et al., 2007), pitch (Dolscheid et al., 2013), odor (Majid, 2015), representation of time (Bylund and Athanasopoulos, 2017; Casasanto, 2008), and numerosity (Pica et al., 2004) (for a review see (Brysbaert, 2018)). These recent investigations have provided renewed interest and generated broad discussions (Casasanto, 2016; Kremmerer, 2019; Malt and Wolff, 2010) on the Sapir-Whorf hypothesis (Sapir, 1929; Whorf, 1956), now reformulated in terms of predictive processing (Lupyan and Clark, 2015) and probabilistic inference (Cibelli et al., 2016; Regier and Xu, 2017). These provide new accounts of how our thoughts can be shaped by the way we speak, particularly in situations of high uncertainty. The label-feedback hypothesis (LFH) (Lupyan, 2012) removes the distinction between verbal and non-verbal processes, with labels providing transient top-down predictive signals sharpening perceptual processes. A range of studies has already shown that object labels can facilitate the identification of objects (Boutonnet et al., 2013; Boutonnet and Lupyan, 2015; Lupyan et al., 2007; Lupyan and Ward, 2013). Here we test and demonstrate that the modulating effects of labels can be extended to influence motor actions, and our ability to learn and use tools. According to embodied theories of cognition (Barsalou, 2010; Clark, 1999; Meteyard et al., 2012), language is grounded in perception, action, and our sensorimotor environment. We propose that, by naming a tool, we create an invariable linguistic representation that provides an anchor for sensorimotor experience, increasing saliency, and helping us build motor learning and expertise.

Here, we trained participants in the use of six novel tools in an immersive Virtual Reality (VR) environment (Fig. 1A). Firstly, participants learned to associate three of these tools with a specific label whereas three other tools were left unnamed (label learning task; Fig. 1B). Secondly, participants learned a functional manipulation unique to each of the six tools (training task; Fig. 1C). Thirdly, the participants engaged in a move-use task, in which they could be asked to use the tool on a target object, or simply to move it to a location. If a label simply helps us identify a tool then it should facilitate faster initiation times when both using and moving a labeled tool (Boutonnet et al., 2013; Boutonnet and Lupyan, 2015; Lupyan et al., 2007; Lupyan and Ward, 2013). If labels help in the learning of the manipulations required to use tools, then participants should be faster to execute an action with a labeled versus unlabeled tool, but not when moving them.

During the move-use task we used EEG to measure event-related synchronization/desynchronization (ERD/ERS) associated with the onset of the presentation of the tools. Analyses focused upon oscillatory activity in the high beta-band (20-40 Hz), well established as an indicator of motor processing (Kilavik et al., 2013; McFarland et al., 2000; Pfurtscheller, 1992;

Pfurtscheller and Lopes, 1999; Turella et al., 2016), as well as action semantics (van Elk et al., 2010a, 2010b), semantic memory (Slotnick et al., 2002) and language processing (Weiss and Mueller, 2012). Neural models propose that ~30 Hz thalamocortical activities reflect the retrieval of semantic information about objects (Hart and Kraut, 2007; Kraut et al., 2003; Slotnick et al., 2002), such as colors or shapes, with beta-band ERD/ERS also found to index the activation of motor knowledge via language (Bechtold et al., 2018). Therefore, we hypothesize that if labels help to facilitate the learning of motor programs associated with a novel tool, this should be reflected in a reduction in ERD/ERS beta-power when participants need to retrieve manipulative information to use a tool, but not when moving it. Conversely, if labels simply facilitate the identification of the tools, then this reduction in beta-power should be present in both tasks.

2. Material and methods

2.1. Participants

Forty adult volunteers (seven males, mean age = 21.2, years old, SD = 6.2) from the University of Plymouth participated for course credit. All participants reported being right-handed and having normal or corrected-to-normal vision. This sample size was chosen to collect an amount of data as similar as the found in the EEG literature. Data from one participant was removed from the analysis due to a technical problem. Protocols were approved by the ethics committee of the University of Plymouth and conform to the 2008 Helsinki Declaration. Our sample size, data exclusions, inclusion/exclusion criteria, inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

2.2. Experimental Design

Participants interacted in a virtual environment viewed through a VR headset (HTC Corp.) and manipulated using a hand-held controller. This environment provided a virtual simulacrum of the physical space; the participants seated at a desk facing a physical push button and a marked space representing the 'home' location for the physical manipulator. The first part of the experiment introduced six novel tools to the participants, each having a unique visual representation, but mapped onto the same physical manipulator (Fig. 1A). Three of these tools were presented with monosyllabic non-word names ("Sni", "Unt" and "Lum"), three were left unnamed. The association of the tools and their names was reinforced and tested by asking participants to move the tools to an appropriately labeled location (Fig. 1B). Tools without names were moved to a location marked with an 'x'. In the functional training phase (Fig. 1C) participants learned to associate each of the tools with a specific function, motor sequence, and target object (e.g. shaking the manipulator and pressing a particular button over a plant to make it grow). EEG was recorded during the final 'move-use' task (Fig. 1D), where participants were cued to either move the tool to a target location next to its associated tool or use the tool on the object using the learned motor sequence. Contrasting the movement and use of objects allows us to ascertain whether labels simply facilitate the identification of the object, or if they help in the learning of motor

processing associated with their function. The required action (Move-Use) was cued with an auditory tone heard at the same time as the onset of the presentation of each tool. Further details on the procedure can be found in the electronic supplementary material, Experimental Tasks. The study procedure and analyses were not pre-registered before the research being conducted.

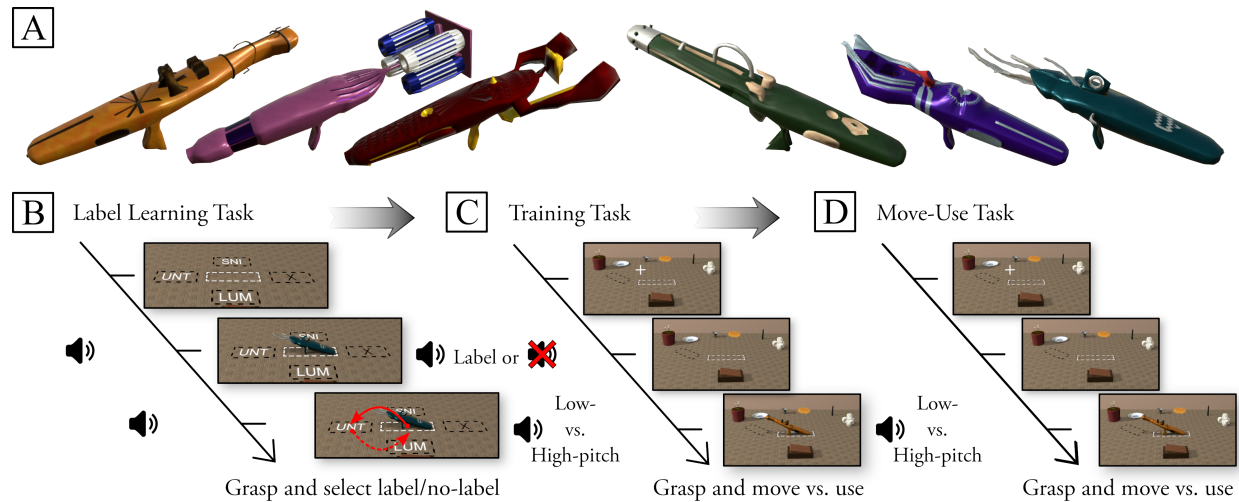


Fig 1. Experimental Design. **A**, Visual representations of the six novel tools manipulable in VR. **B**, Participants learned the labels (“Sni”, “Unt” or “Lum”) of three of the six novel tools in a label learning task. **C**, Participants learned a specific tool use. **D**, EEG and behavioural timings were recorded while participants were instructed to move or use the labeled and unlabeled tools, depending on a high- or low-pitched tone triggered at tool appearance.

2.3. Electrophysiological Recording and Processing

EEG was recorded and sampled at 500 Hz during the move-use task of the experiment from 62 actively amplified Ag/AgCl electrodes (actiCAP, Brain Products, Gilching, Germany) using a BrainAmp MR Plus amplifier (Brain Products). EEGs recordings were analysed with Brain Vision Analyzer (Brain Products, Munich, Germany, v. 2.1) and filtered on-line with a .1 Hz high pass filter, a 50 Hz low pass filter and a 50 Hz notch filter. Electrodes were re-referenced off-line to the average of left and right mastoid activities and the fronto-central electrode AFz was used as the ground. ERPs were time-locked on the visual onset of the presentation of the tool with 1800 ms time window spanning from -600 to 1200 ms before and after the time-lock. Trials with the press button released before the tool onset or within the first 200 ms following the tool onset were discarded. A semi-automatic trial rejection procedure was run on these ERPs to exclude segments violating the following parameters: maximal allowed voltage step of 50 $\mu\text{V}/\text{ms}$, maximal voltage differences allowed of 150 μV within 100 ms intervals, maximal/minimal allowed amplitude of $\pm 120 \mu\text{V}/\text{ms}$, and minimum amplitude of 0.5 μV within 100 ms intervals. These parameters were slightly adapted manually for each participant to maximize the signal/noise ratio and resulted in a total of 15% of segments rejected. Individual electrodes having greater than $\sim 8\%$ of rejected segments were removed from analyses and substituted with topographically interpolated replacements (Perrin et al., 1989), representing a total of $\sim 1\%$ of electrodes replaced over the cohort of participants. Event-Related Desynchronization/Synchronization (ERD/ERS) was calculated for each segment of the ERPs using

a pass-band filter ranging from 20 to 40 Hz. Resulting ERD/ERS amplitude values were then squared and traces were smoothed using a running average time-window of 45 msec, as in Ruther et al. (2014). Each power value was converted in a percentage of power change relative to the average power calculated over a baseline period. This baseline period concerned the immediate 200 ms before the tool and tone onsets and reflected the best segment to control for pre-stimuli noise based on pre-analyses. Finally, each trial was re-segmented to a period 200 ms before the time-lock and 500 ms to remove edge effects inherent in decomposition analyses of segmented data and focus analyses on the temporal period of interest (~ 400 ms) and prior participants' motor response (~ 530 ms). Averaged ERD/ERS traces were calculated separately for each subject and each tool and task (i.e. labeled and move, labeled and use, unlabeled and move, unlabeled and use), then analysed across the whole scalp.

2.4. Statistical analyses

Our tests aimed at the interaction effect of the within-subject variables of labels (labeled vs unlabeled tools) in the motor tasks (move vs use of the tools). Analyses were conducted on behavioural measures taken during the Move-Use phase of the experiment. Measures of accuracy were based upon the participant's ability to apply the presented tool using the correct motor sequence on the appropriate target tool (use) or move the tool to the appropriate location(move). We calculated the percentage of success to perform the tool use for each participant and each trial block, depending on the tool (labeled vs unlabeled). The strategy used for the following modelling was to maximize the complexity of the structure to control for a maximum of variance while keeping converging models given the size of our dataset. In each model the tool (labeled or unlabeled), the trial block (first, second or third) were entered as fixed effects, with the participant as a random effect and by-participant random slopes for the effect of the trial block. P-values were obtained by likelihood ratio tests of the full model against the null model, with and without the interaction term for the tool and the trial block, respectively (formula of the full model: $\text{lmer}(\text{PercentOfSuccess} \sim \text{Tool} + \text{TrialBlock} + \text{Tool} * \text{TrialBlock} + (1|\text{Subject}) + (0 + \text{TrialBlock}|\text{Subject}))$). Subsequent contrasts evaluated the influence of the tool (labeled vs unlabeled) within each trial block, without the interaction and random effect term, and by dropping the tool term for the null model (formula of the full model: $\text{lmer}(\text{PercentOfSuccess} \sim \text{Tool} + \text{TrialBlock} + (1|\text{Subject}))$). Reaction times were also measured at three points during this process: a) Initiation time, the duration between tool presentation onset and button release; b) Grasping time, the duration between button release and the grasp of the tool; c) Execution time, the duration between grasp onset and completion of the move/use action. Concerning the analysis of the reaction times, in each model the tool (labeled or unlabeled), the task (move or use) and the trial block (first, second or third) were entered as fixed effects, with the participant as a random effect and by-participant random slopes for the effect of the trial block. The trial block variable was entered in the models as it represented an important source of variability fluctuating during the experiment. P-values were obtained by likelihood ratio tests of the full model against the null model, with and without the interaction term for the tool and the task, respectively (formula of the full model: $\text{lmer}(\text{ReactionTime} \sim \text{Tool} + \text{Task} + \text{Tool} * \text{Task} + \text{TrialBlock} + (1|\text{Subject}) + (0 + \text{TrialBlock}|\text{Subject}))$). Subsequent contrasts evaluated the influence of the tool (labeled or unlabeled) for each reaction time

and individual task, without the interaction term and by dropping the tool term for the null model (formula of the full model: $\text{lmer}(\text{ReactionTime} \sim \text{Tool} + \text{Task} + \text{TrialBlock} + (1|\text{Subject}) + (0+\text{TrialBlock}|\text{Subject}))$). RStudio (v. 1.1.456) and the lme4 (v. 1.1.1-12) were used to calculate separate linear mixed-effect models analyses for accuracy and each of the three reaction times. Visual inspection of the residuals' plots did not reveal any violation of the assumptions of application. Pseudo-R-squared effect sizes were estimated with the r.squaredLR function of the MuMIn (v. 1.15.6) packages.

Statistical analysis of the ERD/ERS traces opposing the labeled and unlabeled tools within each type of action (move or use) was conducted using a pairwise comparison based on a cluster randomisation technique (Maris and Oostenveld, 2007) to avoid multiple comparisons. Two-tailed t-tests were performed, comparing each electrode-time and electrode-signal sample pair for the labeled and unlabeled tools, separately for each type of action for the whole time-window. Those samples with t statistic above the significance threshold of $p < 0.05$ were clustered together in spatial and temporal terms. Each cluster was based on a minimum of eight samples and used for the subsequent cluster analysis. The cluster-level t statistic was calculated as the sum of the t statistic of all electrode-time samples of a given cluster. For the cluster analysis, the cluster with the largest t statistic was selected for a Monte-Carlo simulation. Thus, each of the original pairs of t-tests sample that compose the cluster was repeated 1000 times, with permutations of each paired samples randomly assigned to the labeled or unlabeled tools for a given type of action. It generated a Monte-Carlo distribution of summed t statistic corresponding to the null hypothesis. A Monte-Carlo p-value was calculated as the ratio of the 1000 summed t statistics in the random distribution that was above the cluster-level t statistic. This p-value was considered significant above $p < 0.025$. Averaged ERD/ERS traces were re-plotted as t-values in the time domain, derived from t-tests against baselines of zero. For a good visualization of the effect, these t-values of the significant cluster had been used to create topographic maps in Brain Vision Analyzer (Brain Products, Munich, Germany, v. 2.1), using spherical spline interpolation with an order of splines of 5 and a maximum degree of Legendre polynomials of 10.

3. Results

3.1. Behavioural Results

Both accuracy and reaction latencies were recorded during the 'move-use' task of the experiment. Trials with initiation times below 200 ms were discarded as erroneous (2.99% of the trials). Two participants displaying inefficient tool use learning were removed from the analysis (the inclusion/exclusion of these participants did not influence the statistical results of the reaction times below). The accuracy reflects participants' ability to recall and execute motor responses appropriate to the utility of the tool. For 'use' trials, a response was marked as accurate if participants were able to perform the correct sequence of tool manipulations adjacent to the correct target object within four seconds from the trial onset tone (93.04% of trials). A mixed-effects model of accuracy with the factors of label and trial block did not reveal a significant difference between labeled (Fig. 2A; $M_{\text{labeled}} = 93.9\%$, $SD = 8.6\%$) and unlabeled tools ($M_{\text{unlabeled}} = 92.8\%$, $SD = 8.2\%$; $\chi^2(3) = 1.64$, $r^2 = 0.007$, $p = 0.2$) nor interaction

effect between the labelling of tools and the trial block ($\chi^2(2) = 1.12, r^2 = 0.005, p = 0.29$). However, the planned comparisons describing the learning effect throughout the experiment revealed that the effect of labelling was not significant in the first block ($M_{\text{labeled}} = 86.2\%, SD = 10.4\%; M_{\text{unlabeled}} = 86.7\%, SD = 10.6\%; \chi^2(2) = 0.06, r^2 < 0.001, p = 0.79$), but significant in the second ($M_{\text{labeled}} = 97.7\%, SD = 3.6\%; M_{\text{unlabeled}} = 95.5\%, SD = 4.6\%; \chi^2(2) = 5.28, r^2 = 0.071, p = 0.021$) and third blocks ($M_{\text{labeled}} = 98.1\%, SD = 3.5\%; M_{\text{unlabeled}} = 96.2\%, SD = 4.1\%; \chi^2(2) = 4.32, r^2 = 0.058, p = 0.037$). For ‘move’ trials, the responses were accurate when the tool was placed next to the correct target object within four seconds from the trial onset tone (97.9% of the trials) and without performing a tool use at first (i.e. a confusion in the meaning of the two auditory cues). The mixed-effects models did not reveal main effect ($p = 0.84$), interaction effect ($p = 0.35$) nor effects in the planned comparisons (all $p > .36$) of the label on the accuracy during these ‘move’ trials.

Three reaction times were analysed using separate mixed-effects models: 1) initiation time, i.e. the latency between the onset-tone and release of the response button, 2) grasping time, i.e. the latency between the initiation and the grasp of the tool, and 3) execution time, i.e. the latency between grasp and completion of the move or use action. These were based on accurate responses with latencies within three standard deviations from each participant’s mean for each individual trial block, task and reaction times (5.01% of accurate trials excluded). Analyses revealed a main effect of the label on initiation times, with shorter RTs for labeled tools than for unlabeled tools ($M_{\text{labeled}} = 521 \text{ ms}, SD = 156 \text{ ms}; M_{\text{unlabeled}} = 527 \text{ ms}, SD = 176 \text{ ms}; \chi^2(2) = 5.73, r^2 < 0.001, p = 0.017$). The analysis of the planned comparisons revealed the effect was only significant within the use task ($M_{\text{labeled}} = 524 \text{ ms}, SD = 155 \text{ ms}; M_{\text{unlabeled}} = 532 \text{ ms}, SD = 199 \text{ ms}; \chi^2(2) = 7.37, r^2 = 0.002, p = 0.007$). There were also significant interactions between the factors of ‘move-use’ and label in grasping ($\chi^2(2) = 4.16, r^2 < 0.001, p = 0.041$) and execution ($\chi^2(2) = 8.45, r^2 = 0.008, p = 0.04$) latencies. Post hoc analyses of these interactions revealed that grasping times were increased when participants were asked to move a labeled tool compared to an unlabeled one ($M_{\text{labeled-move}} = 962 \text{ ms}, SD = 327 \text{ ms}; M_{\text{unlabeled-move}} = 945 \text{ ms}, SD = 289 \text{ ms}; \chi^2(2) = 4.78, r^2 = 0.001, p = 0.028$), while execution times were reduced when using a labeled tool ($M_{\text{labeled-use}} = 1064 \text{ ms}, SD = 627 \text{ ms}; M_{\text{unlabeled-use}} = 1109 \text{ ms}, SD = 668 \text{ ms}; \chi^2(2) = 5.14, r^2 = 0.001, p = 0.023$). Descriptive analysis of the performance to the recall task indicated that 89% of the participants were able to recall the category (labeled vs unlabeled) of the tool, which is selecting a tile with a label when the tool has indeed a label. The correct label was selected 69% of the labeled tools, suggesting an efficient consolidation of the tools’ identity one hour after learning their labels.

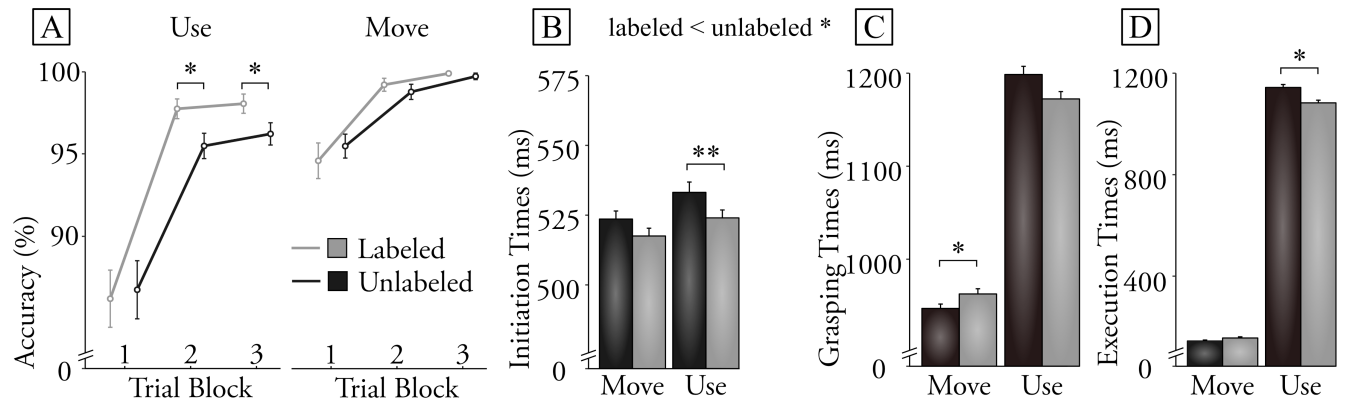


Fig 2. Behavioural results. **A**, Tool use and move performance, reflecting the participants' accuracy to remember how to use and where to move the tools. **B**, Initiation times representing the time necessary to initiate the reach of the tools. **C**, Grasping Times representing the time between the initiation of the reach and the actual grasp of the tools. **D**, Execution Times representing the time necessary to move or use the tools once handled. Error bars show ± 1 SEM (* $p < 0.05$; ** $p < 0.01$).

3.2. EEG Results

At first, relative baseline-corrected ERD/ERS derived from all electrodes between the onset of the visual presentation of tools/action cueing tone and 500 ms post-onset were tested. Separate pairwise comparisons of averaged ERD/ERS for labeled and unlabeled tools were conducted for the tool move (Fig. 3A) and use (Fig. 3B) conditions using the cluster randomization technique (Maris and Oostenveld, 2007). These analyses revealed that, when participants were required to use the tool, the label reduced the beta-band power recorded over a significant ($P < 0.002$) cluster of activity spanning somatosensory and motor areas between 230-500 ms post-onset. The beta-band modulation revealed by the analysis concerned the electrodes Fz, F1, F2, FCz, FC1, FC2, FC3, Cz, C1, C2, C3, C5, CPz, CP1, CP2, CP3, CP4, Pz, P1, P3, P5, P7, POz, PO3, PO4, PO8, Oz, O1, and O2, with a peak of t-values around medial and left-lateralized centro-parietal electrodes at 400 ms post-onset. However, when participants were required to move the tools there were no significant differences in beta-band activity ($p > 0.05$).

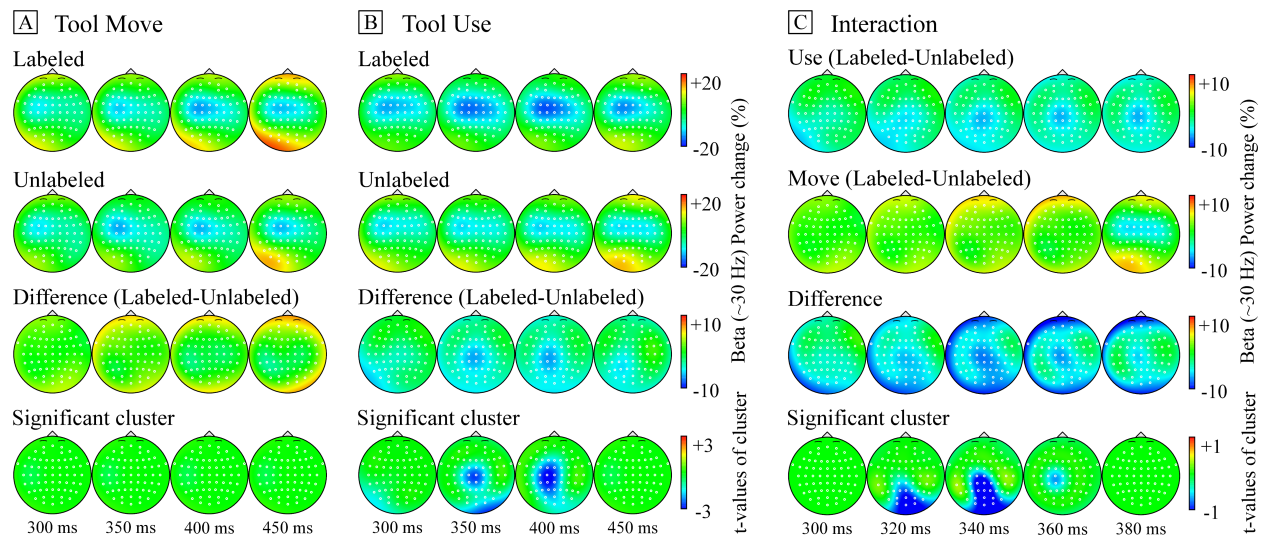


Fig. 3 Topographic maps of the evoked beta-band (20-40 Hz) power at tool onset depending on the labeling of the tool and the task at hand. **A**, Comparison of the labeled and unlabeled tools when required to move. **B**, Comparison of the labeled and unlabeled tools when required to use. Statistical analysis revealed that labeled tools reduced ERD/ERS power only when using an object, not when moving it. This reduction was significant over left sensorimotor and parietal areas and peaking at around 350-400 ms from tool perception. **C**, Comparison of the effect of the label within each motor task calculated as $(\text{Use}_{\text{labeled}} - \text{Use}_{\text{unlabeled}})$ vs $(\text{Move}_{\text{labeled}} - \text{Move}_{\text{unlabeled}})$. Statistical analysis validated the interaction effect, such as labeled tools induced beta decrease during the preparation of tool use only. The cluster of signal decrease appears over centro-parieto-occipital areas at around 340 ms from tool perception.

To validate the interaction effect, we also statistically compared the difference of signal induced by the labels within tool use and move, using the following comparison: $(\text{Use}_{\text{labeled}} - \text{Use}_{\text{unlabeled}})$ vs $(\text{Move}_{\text{labeled}} - \text{Move}_{\text{unlabeled}})$. Given the time-window of the effect to test, the analysis only concerned the ERD/ERS traces from all electrodes between 200 and 500 ms post-onset. The single significant cluster (Fig. 3C; single cluster ranging from 304 to 360 ms, $p = 0.013$) revealed a decrease of beta-band power only when intended to use the tool. The significant modulation concerned the electrodes C1, Pz, P2, CPz, CP1, POz, Oz, and O2. This indicates that the sensorimotor power reduction in the ~30 Hz frequencies due to the labeling of tools is only present when participants need to retrieve information on the use of the tool, rather than simply identify it, as is the case in the move condition.

4. Discussion

What are the consequences of knowing the name of the tools we are learning to use? Our experimental data extend the view proposed by Sapir (Sapir, 1929) and Whorf (Whorf, 1956) showing that linguistic labels influence not only the way we think, but also the way we act.

We found that our participants were faster both to grasp and to use novel tools that were labeled, compared to unnamed tools. This could simply reflect a better identification of the tools and provide further support to the LFH (Lupyan, 2012; Lupyan and Clark, 2015; Lupyan and Ward, 2013). Crucially, however, adopting the move-use task allowed us to distinguish the effect of the labels on perceptual and/or action systems. We hypothesized that tool use would rely more on motor learning and would benefit more from top-down sensorimotor beta-band signals driven by lexical representations. This suggests that labels strengthen the association of stable properties (e.g. functional grip to use the tools) with their referent during learning. In accordance, our EEG data revealed that the benefit of the labels is at least partially action-goal specific. When intending to use the tools, labeled tools induced a decrease of ~30 Hz beta-band power over somatosensory and motor areas. The location of the modulation suggests that adding a lexical representation to a novel tool guides its grounding into specialized action systems. Below, we propose a possible mechanism for such label-augmented motor learning.

The two action systems theory (2AS) (Buxbaum and Kalenine, 2010) and its updated version (2AS+) (Buxbaum, 2017), proposes that the use of tools requires a left-lateralized 'Function' system storing and retrieving action representations whereas a bilateral 'Structure' system processes perceived geometries of objects allowing reach and grasp actions. During the preparation of tool manipulation, the two systems are activated and may interfere with each other. This resonates with the affordance competition hypothesis (Cisek, 2007; Cisek and Kalaska, 2010), suggesting that action selection ('what') and specification ('how') of movements are continuously competing before and during the action. Non-motoric information would interfere with this competition between the multiple manipulations available and bias the selection of the most appropriate behaviour. Our data indicate that learning object labels affect the selection of handgrips and could even foster skilled tool use implemented by the Function system. Thus, beyond perception, the use of language can shape our manual skills.

Why would learning the label of a tool influence our ability to manipulate it? According to the connectionist model of the LFH, labels associated with an object are automatically activated when it is viewed. These invoke top-down regulation of perceptual processes to guide perceptual experience towards visual details most relevant to recognition or categorization. In the perceptual domain, labels help to generate predictions on noisy visual inputs and render visual discrimination more effective (Lupyan and Ward, 2013). In the motor domain, the 2AS+ theory proposes that manipulation knowledge of objects generates motor predictions about desired body states, while sensory predictions are used to minimize the error in motor commands (Buxbaum, 2017; Buxbaum and Kalenine, 2010). Practicing uncertain tool use would refine predictions of learned body states, effectively what we expect to see and feel when using a tool. When labeled, visual predictions generated during the use of the tool would facilitate the learning of these manipulations. Because neural processing involves highly interactive feed-forward and feedback loops, the benefit of learning a label would rely on the interaction between multiple hierarchical levels of information processing. A recent study reported a label-augmented discrimination skill of vibrotactile stimuli (Miller et al., 2018), where learning the label of Braille-like stimuli increased the connectivity strength between hippocampus, auditory and somatosensory cortical regions (Schmidt et al., 2019). In accordance with our results, an fMRI study reported that associating novel names with novel knots affects the ability to discriminate these knots and increases parietal activities (Cross et al., 2012). Importantly, the beta desynchronization recorded over sensorimotor areas has been associated with lexical processing (Grabner et al., 2007). In this study, the processing of words induced an increase of beta desynchronization at around 400 ms post-stimulus onset when compared to pseudo-words. Our beta-band effect could reflect an augmented learning of tool use states within the somatosensory regions via the (re)activation of the lexical representation. Both the timing and topography of the cluster reflecting the beta-band effect support this interpretation. An alternative explanation based upon perceptual facilitation would be more consistent with earlier occipital activation, within 100 ms of object recognition (Boutonnet and Lupyan, 2015).

The similarities between our findings and those of prior studies are consistent with our hypothesis that labels can influence the learning of tool use, based on a mechanism similar to that espoused for visual affordances. Affordances represent the practical opportunities for interaction (Gibson, 1979) in transforming percepts into motor information, such as the visualization of a tool would prime a grasping by its handle. A variety of studies have shown that visual objects automatically evoke affordances (Goslin et al., 2012; Symes et al., 2005; Tucker and Ellis, 1998; Vainio et al., 2008). Visual affordances can be highly variable, reflecting a temporary relationship with the object (e.g. orientation and distance relative to the viewer). Other affordances can be abstracted over experience and encoded in stable object representations (e.g. utility). Concerning affordances an important question remains. Given that the visual processing of the object might be altered by language (Boutonnet and Lupyan, 2015; Lupyan and Ward, 2013), is the extraction of the affordances preserved? Indeed, specific affordances might be bolstered despite others. The present data cannot support the idea of altered bottom-up affordances processing. However, the competition between visual and learned affordances occurring within the fronto-parietal action selection network might be biased by the presence of a label, in favorizing the processing of tool-use memories over geometrical properties. Another possibility is that our lexicon extends motor learning capabilities, such as the presence of a label reinforces the abstraction of recurrent information relevant for action. This improved abstraction would explain the increased ceiling performance of usage for the labeled tool. In accordance with the present beta modulation, stable affordances would be represented in parietal ventro-dorsal circuits (involving the anterior supramarginal gyrus and human putative anterior intraparietal area) (Orban and Caruana, 2014; Sakreida et al., 2016) rather than bilaterally. Our data support a recent proposal on the role of language in processing stable/variable affordances (Borghi, 2012; Borghi and Riggio, 2015, 2009) suggesting that language filters and encodes specifically stable tool properties.

5. Conclusion

We propose that, by encoding stable tool properties, language can help us acquire the functional and motor properties of tools. In providing a novel tool with a name, we ground uncertain tool use experiences around an invariant linguistic representation, helping our acquisition of the motor skill. Here, we bring behavioural and neurophysiological evidence that the top-down activation provided by linguistic labels not only helps to ground conceptual information, as proposed by the LFH (Lupyan, 2012) but also plays a role in motor learning.

Addressing the Whorfian question of whether linguistic and non-linguistic processes are cognitively distinct, recent studies have adopted a view that language and perception are co-dependent (Athanasopoulos et al., 2009; Bylund and Athanasopoulos, 2017; Dolscheid et al., 2013; Winawer et al., 2007). Here we presented evidence to suggest that this dependency is not restricted to perception, but also extends to the motor system. Naming a novel tool appears to help us learn how to use it,

providing a “neuroenhancement” (Dove, 2018) linking lexical, semantic and motor brain representations. The lexical representation providing an anchor for sensorimotor experience, helping to guide the acquisition, selection and recall of motor programs associated with tool manipulation during learning.

Acknowledgements

This work was supported by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement [642667]. Behavioural data and R scripts are available in <https://github.com/FrancoisFoerster/Labels-strengthen-motor-learning-of-new-tools/>. An online video describing the experiment is accessible via <https://www.youtube.com/watch?v=5nowjs1dO8c>.

References

- Arbib, M.A., 2011. From Mirror Neurons to Complex Imitation in the Evolution of Language and Tool Use. *Annu. Rev. Anthropol.* 40, 257–273. <https://doi.org/10.1146/annurev-anthro-081309-145722>
- Athanasopoulos, P., Wiggett, A., Dering, B., Kuipers, J.R., Thierry, G., 2009. The Whorfian mind: Electrophysiological evidence that language shapes perception. *Commun. Integr. Biol.* 2, 332–334. <https://doi.org/10.4161/cib.2.4.8400>
- Barsalou, L.W., 2010. Grounded Cognition: Past, Present, and Future. *Top. Cogn. Sci.* 2, 716–724. <https://doi.org/10.1111/j.1756-8765.2010.01115.x>
- Bechtold, L., Ghio, M., Lange, J., Bellebaum, C., 2018. Event-related desynchronization of mu and beta oscillations during the processing of novel tool names. *Brain Lang.* 177–178, 44–55. <https://doi.org/10.1016/j.bandl.2018.01.004>
- Borghi, A.M., 2012. Action language comprehension, affordances and goals, in: Coella, Y., Bartolo, A. (Eds.), *Language and Action in Cognitive Neuroscience. Contemporary Topics in Cognitive Neuroscience Series*. London: Psychology Press, pp. 125–143.
- Borghi, A.M., Riggio, L., 2015. Stable and variable affordances are both automatic and flexible. *Front. Hum. Neurosci.* 9, 351. <https://doi.org/10.3389/fnhum.2015.00351>
- Borghi, A.M., Riggio, L., 2009. Sentence comprehension and simulation of object temporary, canonical and stable affordances. *Brain Res.* 1253, 117–128. <https://doi.org/10.1016/j.brainres.2008.11.064>
- Boutonnet, B., Dering, B., Viñas-Guasch, N., Thierry, G., 2013. Seeing Objects through the Language Glass. *J. Cogn. Neurosci.* 25, 1702–1710. https://doi.org/10.1162/jocn_a_00415
- Boutonnet, B., Lupyan, G., 2015. Words Jump-Start Vision: A Label Advantage in Object Recognition. *J. Neurosci.* 35, 9329–9335. <https://doi.org/10.1523/JNEUROSCI.5111-14.2015>

- Brysbaert, M., 2018. Numbers and language: What's new in the past 25 years?, in: Henik, A., Fias, W. (Eds.), *Heterogeneity of Function in Numerical Cognition*. Elsevier: Academic Press, pp. 3–26.
- Buxbaum, L.J., 2017. Learning, remembering, and predicting how to use tools: Distributed neurocognitive mechanisms: Comment on osiurak and badets (2016). *Psychol. Rev.* 124, 346–360. <https://doi.org/10.1037/rev0000051>
- Buxbaum, L.J., Kalenine, S., 2010. Action knowledge, visuomotor activation, and embodiment in the two action systems. *Ann. N. Y. Acad. Sci.* 1191, 201–218. <https://doi.org/10.1111/j.1749-6632.2010.05447.x>
- Bylund, E., Athanasopoulos, P., 2017. The Whorfian time warp: Representing duration through the language hourglass. *J. Exp. Psychol. Gen.* 146, 911–916. <https://doi.org/10.1037/xge0000314>
- Casasanto, D., 2016. Linguistic Relativity, in: Riemer, N. (Ed.), *Routledge Handbook of Semantics*. New York: Routledge, pp. 158–174. <https://doi.org/10.1515/9783110308143>
- Casasanto, D., 2008. Who 's Afraid of the Big Bad Whorf ? Crosslinguistic Differences in Temporal Language and Thought. *Lang. Learn.* 58, 63–79.
- Cibelli, E., Xu, Y., Austerweil, J.L., Griffiths, T.L., Regier, T., 2016. The Sapir-Whorf Hypothesis and Probabilistic Inference: Evidence from the Domain of Color. *PLoS One* 11, 1–28. <https://doi.org/10.1371/journal.pone.0158725>
- Cisek, P., 2007. Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 362, 1585–1599. <https://doi.org/10.1098/rstb.2007.2054>
- Cisek, P., Kalaska, J.F., 2010. Neural Mechanisms for Interacting with a World Full of Action Choices. *Annu. Rev. Neurosci.* 33, 269–298. <https://doi.org/10.1146/annurev.neuro.051508.135409>
- Clark, A., 1999. An embodied cognitive science? *Trends Cogn. Sci.* 3, 345–351. [https://doi.org/10.1016/S1364-6613\(99\)01361-3](https://doi.org/10.1016/S1364-6613(99)01361-3)
- Cross, E.S., Cohen, N.R., Hamilton, A., Ramsey, R., Wolford, G., Grafton, S.T., 2012. Physical experience leads to enhanced object perception in parietal cortex: Insights from knot tying. *Neuropsychologia* 50, 3207–3217. <https://doi.org/10.1016/j.neuropsychologia.2012.09.028>
- Dolscheid, S., Shayan, S., Majid, A., Casasanto, D., 2013. The Thickness of Musical Pitch: Psychophysical Evidence for Linguistic Relativity. *Psychol. Sci.* 24, 613–621. <https://doi.org/10.1177/0956797612457374>
- Dove, G., 2018. Language as a disruptive technology: Abstract concepts, embodiment and the flexible mind. *Philos. Trans. R. Soc. B Biol. Sci.* 373. <https://doi.org/10.1098/rstb.2017.0135>
- Frey, S.H., 2008. Tool use, communicative gesture and cerebral asymmetries in the modern human brain. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 1951–1957. <https://doi.org/10.1098/rstb.2008.0008>
- Gibson, J.J., 1979. *The Ecological Approach to Visual Perception: Classic Edition*. <https://doi.org/10.1002/bs.3830260313>

- Goslin, J., Dixon, T., Fischer, M.H., Cangelosi, A., Ellis, R., 2012. Electrophysiological examination of embodiment in vision and action. *Psychol. Sci.* 23, 152–7. <https://doi.org/10.1177/0956797611429578>
- Grabner, R.H., Brunner, C., Leeb, R., Neuper, C., Pfurtscheller, G., 2007. Event-related EEG theta and alpha band oscillatory responses during language translation. *Brain Res. Bull.* 72, 57–65. <https://doi.org/10.1016/j.brainresbull.2007.01.001>
- Hart, J.J., Kraut, M.A., 2007. *Neural Basis of Semantic Memory*. Cambridge University Press.
- Kilavik, B.E., Zaepffel, M., Brovelli, A., MacKay, W.A., Riehle, A., 2013. The ups and downs of beta oscillations in sensorimotor cortex. *Exp. Neurol.* 245, 15–26. <https://doi.org/10.1016/j.expneurol.2012.09.014>
- Kraut, M.A., Calhoun, V., Pitcock, J.A., Cusick, C., Hart, J.J., 2003. Neural hybrid model of semantic object memory: Implications from event-related timing using fMRI. *J. Int. Neuropsychol. Soc.* 9, 1031–1040. <https://doi.org/10.1017/S135561770397007X>
- Kremmerer, D., 2019. *Concepts in the Brain: the View from Cross-Linguistic Diversity*. Oxford University Press.
- Lupyan, G., 2012. Linguistically modulated perception and cognition: The label-feedback hypothesis. *Front. Psychol.* 3, 1–13. <https://doi.org/10.3389/fpsyg.2012.00054>
- Lupyan, G., Clark, A., 2015. Words and the World: Predictive Coding and the Language-Perception-Cognition Interface. *Curr. Dir. Psychol. Sci.* 24, 279–284. <https://doi.org/10.1177/0963721415570732>
- Lupyan, G., Rakison, D.H., McClelland, J.L., 2007. Language Is Not Just for Talking. *Psychol. Sci.* 18, 1077–1083. <https://doi.org/10.1111/j.1467-9280.2007.02028.x>
- Lupyan, G., Ward, E.J., 2013. Language can boost otherwise unseen objects into visual awareness. *Proc. Natl. Acad. Sci. U. S. A.* 110, 14196–14201. <https://doi.org/10.1073/pnas.1303312110>
- Majid, A., 2015. Cultural Factors Shape Olfactory Language. *Trends Cogn. Sci.* 19, 629–630. <https://doi.org/10.1016/j.tics.2015.06.009>
- Malt, C.B., Wolff, P., 2010. *Words and the Mind: How words capture human experience*. Oxford University Press.
- Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>
- McFarland, D.J., Miner, L. a, Vaughan, T.M., Wolpaw, J.R., 2000. Mu and beta rhythm topographies during motor imagery and actual movements. *Brain Topogr.* 12, 177–186. <https://doi.org/10.1007/s00221-009-2062-4>
- Meteyard, L., Cuadrado, S.R., Bahrami, B., Vigliocco, G., 2012. Coming of age: A review of embodiment and the neuroscience of semantics. *Cortex* 48, 788–804. <https://doi.org/10.1016/j.cortex.2010.11.002>
- Miller, T.M., Schmidt, T.T., Pulvermüller, F., Blankenburg, F., 2018. Verbal labels facilitate tactile perception. *Cognition* 171, 172–

179. <https://doi.org/10.1016/j.cognition.2017.10.010>
- Orban, G.A., Caruana, F., 2014. The neural basis of human tool use. *Front. Psychol.* 5, 1–12. <https://doi.org/10.3389/fpsyg.2014.00310>
- Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72, 184–187. [https://doi.org/10.1016/0013-4694\(89\)90180-6](https://doi.org/10.1016/0013-4694(89)90180-6)
- Pfurtscheller, G., 1992. Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest. *Electroencephalogr. Clin. Neurophysiol.* 83, 62–69. [https://doi.org/10.1016/0013-4694\(92\)90133-3](https://doi.org/10.1016/0013-4694(92)90133-3)
- Pfurtscheller, G., Lopes, F.H., 1999. Event-related EEG / MEG synchronization and desynchronization : basic principles. *Clin. Neurophysiol.* 110, 1842–1857.
- Pica, P., Lemer, C., Izard, V., Dehaene, S., 2004. Exact and approximate arithmetic in an Amazonian indigene group. *Science* (80-.). 306, 499–503. <https://doi.org/10.1126/science.1102085>
- Regier, T., Kay, P., 2009. Language, thought, and color: Whorf was half right. *Trends Cogn. Sci.* 13, 439–446. <https://doi.org/10.1016/j.tics.2009.07.001>
- Regier, T., Xu, Y., 2017. The Sapir-Whorf hypothesis and inference under uncertainty. *Wiley Interdiscip. Rev. Cogn. Sci.* 8, e1440. <https://doi.org/10.1002/wcs.1440>
- Ruther, N.N., Brown, E.C., Klepp, A., Bellebaum, C., 2014. Observed manipulation of novel tools leads to mu rhythm suppression over sensory-motor cortices. *Behav. Brain Res.* 261, 328–335. <https://doi.org/10.1016/j.bbr.2013.12.033>
- Sakreida, K., Effmert, I., Thill, S., Menz, M.M., Jirak, D., Eickhoff, C.R., Ziemke, T., Eickhoff, S.B., Borghi, A.M., Binkofski, F., 2016. Affordance processing in segregated parieto-frontal dorsal stream sub-pathways. *Neurosci. Biobehav. Rev.* 69, 89–112. <https://doi.org/10.1016/j.neubiorev.2016.07.032>
- Sapir, E., 1929. The Status of Linguistics as a Science. *Linguist. Soc. Am.* 5, 207–214.
- Schmidt, T.T., Miller, T.M., Blankenburg, F., Pulvermüller, F., 2019. Neuronal correlates of label facilitated tactile perception. *Sci. Rep.* 9, 1–8. <https://doi.org/s41598-018-37877-w>
- Slotnick, S.D., Moo, L.R., Kraut, M.A., Lesser, R.P., Hart, J.J., 2002. Interactions between thalamic and cortical rhythms during semantic memory recall in human. *Proc. Natl. Acad. Sci.* 99, 6440–6443. <https://doi.org/10.1073/pnas.092514899>
- Symes, E., Ellis, R., Tucker, M., 2005. Dissociating object-based and space-based affordances. *Vis. cogn.* 12, 1337–1361. <https://doi.org/10.1080/13506280444000445>
- Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., Kuipers, J., 2009. Unconscious effects of language-specific terminology on preattentive color perception. *Proc. Natl. Acad. Sci.* 106, 4567–4570.

- Tucker, M., Ellis, R., 1998. On the relations between seen objects and components of potential actions. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 830–846. <https://doi.org/10.1037/0096-1523.24.3.830>
- Turella, L., Tucciarelli, R., Oosterhof, N.N., Weisz, N., Rumiati, R., Lingnau, A., 2016. Beta band modulations underlie action representations for movement planning. *Neuroimage* 136, 197–207. <https://doi.org/10.1016/j.neuroimage.2016.05.027>
- Vainio, L., Symes, E., Ellis, R., Tucker, M., Ottoboni, G., 2008. On the relations between action planning, object identification, and motor representations of observed actions and objects. *Cognition* 108, 444–465. <https://doi.org/10.1016/j.cognition.2008.03.007>
- van Elk, M., van Schie, H.T., van den Heuvel, R., Bekkering, H., 2010a. Semantics in the motor system: motor-cortical Beta oscillations reflect semantic knowledge of end-postures for object use. *Front. Hum. Neurosci.* 4, 1–12. <https://doi.org/10.3389/neuro.09.008.2010>
- van Elk, M., van Schie, H.T., Zwaan, R.A., Bekkering, H., 2010b. The functional role of motor activation in language processing: Motor cortical oscillations support lexical-semantic retrieval. *Neuroimage* 50, 665–677. <https://doi.org/10.1016/j.neuroimage.2009.12.123>
- Weiss, S., Mueller, H.M., 2012. “Too many betas do not spoil the broth”: The role of beta brain oscillations in language processing. *Front. Psychol.* 3, 1–15. <https://doi.org/10.3389/fpsyg.2012.00201>
- Whorf, B.L., 1956. *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*. Massachusetts Institute of Technology: Cambridge.
- Winawer, J., Witthoft, N., Frank, M.C., Wu, L., Wade, A.R., Boroditsky, L., 2007. Russian blues reveal effects of language on color discrimination. *Proc. Natl. Acad. Sci.* 104, 7780–7785.

Electronic Supplementary Information

1. Experimental Tasks

1.1. Label learning Task

The instructions to the task were displayed on the TV screen located in front of the participant in VR. On the left, right, upper and lower sides of the central white dashed rectangle were placed four black dashed rectangles with the labels “Sni”, “Unt”, “Lum” and “X”. At the beginning of each trial, the locations of the four black rectangles were randomly permuted and one of the six tools (Fig. 14) appeared at the location of the physical VR controller over the central white rectangle. In one possible case, when the tool appeared the participant heard a voice from a speaker in the room, labeling the tool with one of the three possible names. The task of the participant was to grasp and transport the tool on the black rectangle associated with its label. If the tool was placed on the correct rectangle, the same voice was triggered again to reinforce the learning. If the tool was placed on a wrong rectangle, a buzzing sound followed by the voice providing the correct label of the tool was triggered, informing the participant of the wrong name label learning of the tool. Importantly, at the appearance of the tool, the voice was triggered only for the first exposure, forcing the participant to actively remember the name of the tool during the following trials. In the other possible case, no voice was triggered at the appearance of the tool, suggesting the tool doesn’t have a name. In this situation, the participant had to grasp and transport the tool to the black rectangle marked with the “X”. If the tool was not placed on the “X” rectangle, the buzzing sound was triggered. In both cases, once the tool was transported on a black rectangle the participant had to put the VR controller back on the central white rectangle, which initialized the next trial. As a result of the label learning task, a first tool was associated with the voice and label “Sni”, a second with “Unt”, a third with “Lum” (forming the tool category Labeled) and the three other tools were associated with the “X” (forming the tool category Unlabeled). Three sets of voices (two men and one female) for each name were recorded and randomly assigned to each trial. We used multiple voices to strengthen the semantic processing of the voices (i.e. the label) while minimizing the processing of their perceptual properties (e.g. pitch). The participants performed the label learning task until they reach the following requirements: 1) minimum of five trials to each tool, and 2) four correct categorizations over the last five attempts for each tool. These requirements were implemented to assure the learning of the name of each tool before initiating the training task. This label learning task lasted approximately 8-10 minutes for each participant.

1.2. Training Task

In the training task, participants learned to perform a unique tool use with each tool and to associate each tone (low- or high-pitch go-signal) to a type of action (move or use). At the beginning of the training task, the black rectangles of the label learning

task disappeared and six other target-tools plus a single black rectangle appeared on the table. These target-tools were a red container with a plant, a white and blue small plate, a grey cup-like shape, a yellow waffle, a black spike, and four ice cubes. At the beginning of each trial, the locations of these six target-tools were randomly permuted and the location of the black rectangle was randomly assigned in the front of one of these target-tools. Each tool was manipulated seven times in a row, including six trials to use the tool followed by another single trial to move the tool. In total, the training task resulted in 42 trials. For each tool use, on the TV screen were displayed a video of the physical manipulation of the controller (recorded off-line and performed by the experimenter) and a schematic representation of the controller with the different buttons to press. These buttons were a big “pad” button located on the top of the controller, a small “top” button on the top, two “grip” buttons on the sides and a “trigger” button on the back. Instructions were given orally by the experimenter rather than written for a maximum of clarity. During this training task, each tool was associated with a novel function, unique manipulation and paired with one of the six target-tools on the table. We designed novel functions in order to control for the inference of a function, manipulation or target-tool pair from the structural properties the tool (e.g. the red tool could be a weird claw to lift the plant’s container). Each tool was associated with one of the six following tool use: 1) pressing three times a grip button to grow the plant, 2) swinging successively the tool to the right, left and right side to dissolve the plate in dust, 3) holding down the top button for 2 sec to create a rock on the top of the grey cup, 4) pressing the trigger button then rotating the wrist 45° to the left to remove the gravity of the waffle (resulting in a floating effect), 5) pressing successively the trigger, pad and trigger buttons to generate a black hole on the tip of the black spike, and 6) swinging vertically the tool up, down and up again to produce a grey cloud and falling snowflakes on the top of the ice cubes. Importantly, the tools needed to be in close distance from and pointing at the correct target-tools to trigger the specific sound effect and visual animation associated with each tool use. During this training task, black panels located over the target-tools instructed the different manipulations to perform with each tool (**Fig. S1**). To initiate a trial, the participant was required to press and hold down the physical button on the table. Pressing the button resulted in the appearance of a white fixation cross over the white rectangle for 1000 ms before disappearing. If the tool assigned to the trial was associated with a label, this label was spoken by one of the three possible voices. Between 800 ms and 1100 ms after the offset of the fixation cross, one of the two tones was triggered and the virtual representation of the tool appeared. At this moment, the participant had to release the press button, grasp and use the tool with the correct target-tool or move the tool to the black dashed rectangle. Once the action was executed, placing the controller back on the white dashed rectangle initiated the next trial and the visual representation of the tool disappeared. The pairing of the tones to the move-use actions and visual representation of tool to the tool use were completely counterbalanced for each participant. This training task lasted approximately 20 minutes for each participant.

1.3. Move-Use Task

The Move-Use task was very similar to the training task. However, no voices were spoken and the tool and the tone (instructing to move or use the tool) were randomly assigned at the beginning of each trial. Also, black panels and representative tools disappeared from the table. The task of the participant was to perform the use or move manipulation of the tools depending on the tone at tool appearance. If the press button was released before the tool onset or below 200 ms after tool onset, a panel appeared reminding the participant to hold the press button down until she/he decided of the tool manipulation to do. Also, if the participants were not able to perform the move or tool use manipulation in the coming 4 sec following the tool onset, a panel displaying the manipulation to do appeared and lasted until the end of the trial. In both these cases, the trial was considered as failed. At first, participants practiced 20 trials in the presence of the experimenter. During these 20 trials, the chance of the tone representing the tool use actions to appear was of 80% and the tone representing the tool move actions of 20%. We used this bias to train particularly the participants to the most difficult task of remembering how to use the tools. Once the first 20 trials were executed, the participants perform three blocks of 100 trials each. During these blocks, each type of move-use trials had a 50% chance of occurrence and we recorded participants' EEG. At the beginning of the first block and between each block, break times were proposed to the participants in order to remove the VR headset. This move-use task lasted approximately 40 minutes for each participant.

1.4. Recall Task

Before to finish the experiment, we asked participants to perform a recall task in order to evaluate the long-term association of the tool with its name (or no-name). To do so, the procedure was the same as the label learning task, except the following points: 1) no voice or buzz sound was triggered (i.e. no feedback about the correctness or incorrectness of the categorization), and 2) each tool was presented a single time. Hence, the recall task consisted of 6 trials.



Fig. S1. Screenshot of the virtual environment taken at the end of the training task. Black panels with the manipulation instructions were displayed along the tools, the labels were displayed on red panels. Before to start the move-use task, tools and black and red panels disappeared.